طرح پيشنهادي براي  
فاز نخست پروژه بازسازي سامانه هوشمند جُهد؛  
روزآوري بانك‌هاي اطلاعاتي

14 مرداد 1402

سامانه جُهد يكي از ابزارهاي دستيار فقيه است كه با استفاده از هوش مصنوعي و دسترسي به صدها كتاب فقهي، حديثي و لغت از منابع اسلامي؛ شيعه و سني، ياريگر پژوهش‌هاي اجتهادي‌ست؛ در حوزه‌هاي مختلفي مانند: كلام، فقه و اخلاق. اين ابزار پس از مدّتي بررسي و ارزيابي به نظر مي‌رسد نيازمند اصلاح و روزآوري بانك‌هاي داده خود است، به نحوي كه بهتر و دقيق‌تر بتواند پاسخگو باشد.

# وضعيت فعلي سامانه جُهد

در نخستين نگاه اين ابزار مانند بسياري از سامانه‌هاي جستجو صرفاً يك جعبه براي ثبت واژه يا عبارت مورد نظر كاربر دارد.



پس از جستجو نيز نتايج در سه ستون به كاربر نشان داده مي‌شود. ستون راست حاوي آمارهايي از داده‌هاي موجود، ستون مياني يافته‌هاي جُهد در متن كتاب‌هاي موجود و ستون چپ هم حاوي يافته‌هايي از سامانه ديگري به نام «قم‌نت» كه مفاهيم مرتبط را نمايش مي‌دهد.



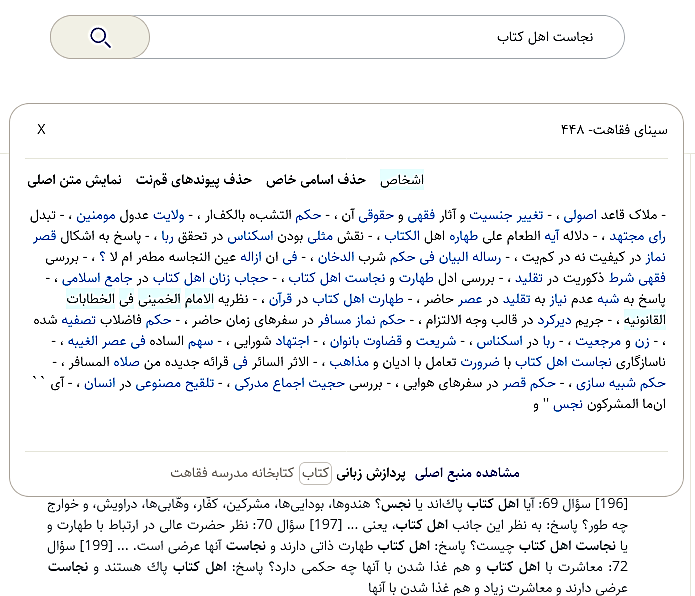
در پايين صفحه نيز جستجوهايي را نشان مي‌دهد كه مرتبط با عبارت ثبت شده توسط كاربر است:



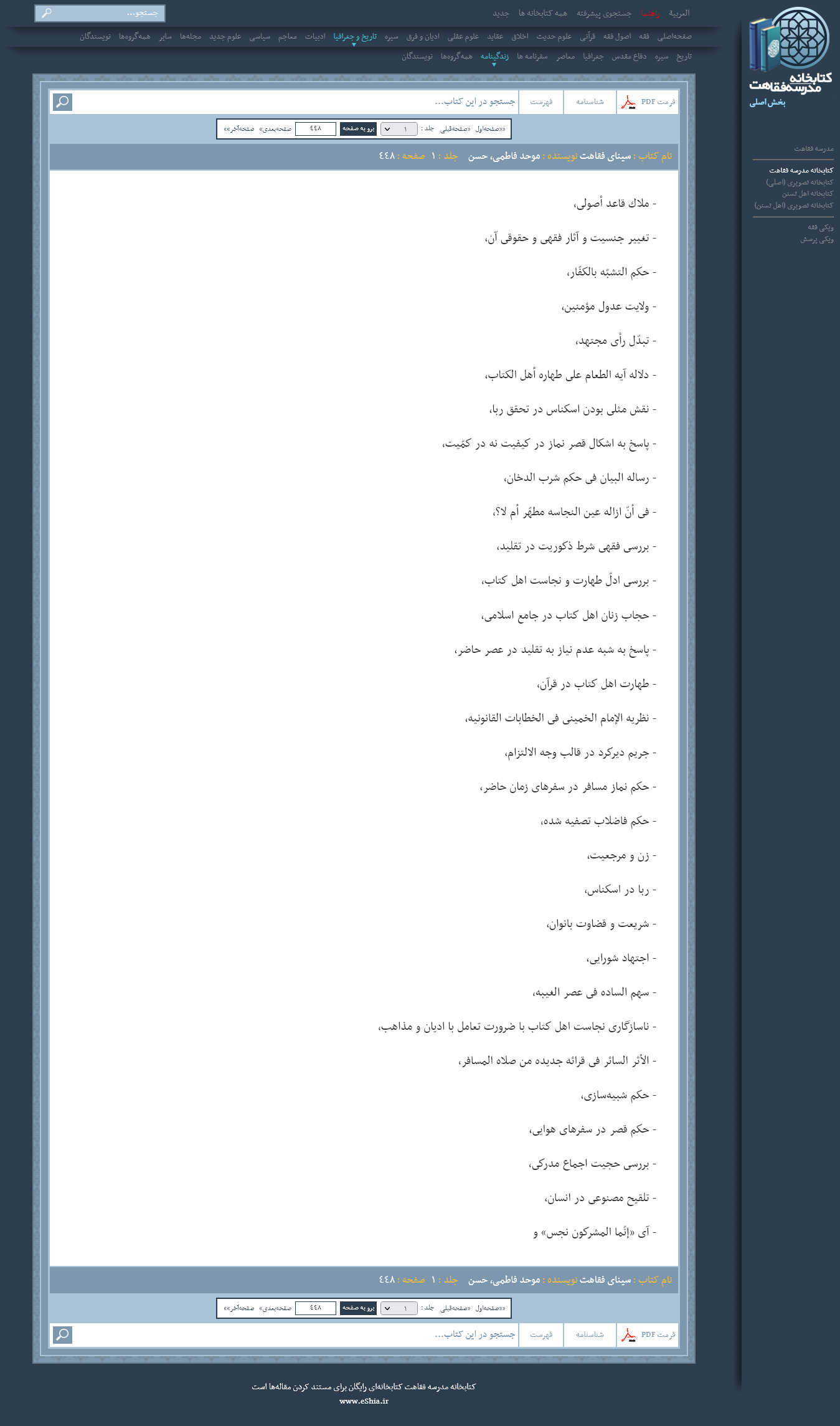
با كليك روي هر كدام از گزينه‌هاي نتايج، پنجره‌اي روي اطلاعات صفحه باز مي‌شود كه همان نتيجه را به صورت تفصيلي‌تر نمايش مي‌دهد:



گزينه پردازش زباني متن يافته شده را با سامانه «قم‌نت» سنجيده و واژگان موجود در آن سامانه را در همان پنجره رنگي كرده و به سامانه اصلي «قم‌نت» لينك مي‌نمايد.



گزينه «مشاهده منبع اصلي»‌ هم فقط يك لينك است به سايتي كه اطلاعات از آن دريافت شده است:



اين تمام چيزي‌ست كه در نگاه اول و پيش از «ورود» به سامانه ديده مي‌شود.

# انتظار پژوهشگر از سامانه جُهد

آن‏چه يك كاربر از اين سامانه انتظار دارد، پژوهشگري كه به دنبال يافتن منابع مورد نياز خود در پژوهش است خيلي فراتر از اطلاعاتي‌ست كه اكنون دريافت مي‌كند:

1. در پايين‌ترين سطح:
   1. فهرست كتاب‌هايي را نشانم بده كه اين موضوع در آ‌ن‌ها مورد بحث قرار گرفته است.
      1. شماره صفحه آن كتاب‌ها را هم در اختيارم بگذار.
      2. همراه با لينك به متن اصلي كتاب‌ها در سامانه منبع.
      3. به من بگو اعتبار علمي كدام كتاب بالاتر و كدام پايين‌تر است.
      4. نام نويسنده كتاب هم در كنارش باشد.
   2. كتاب‌ها را هم از هم تفكيك و دسته‌بندي كن، به صورت موضوعي.
      1. برخي از منابع براي من جنبه لغت‌شناسي دارند؛ كتاب‌هاي لغت.
      2. برخي فقهي هستند و مباحث تخصّصي.
      3. كتاب‌هايي فقط فتوا هستند، مثل رساله‌هاي عمليه.
      4. كتب حديثي هم كاملاً ماهيت متفاوتي دارند. آن‌هايي كه فقط روايات را ذكر كرده‌اند.
   3. دنبال واژه‌ها نيستم. موضوع من يك «عبارت» است.
      1. اگر نوشتم «نجاست اهل كتاب» نمي‌خواهم منابع «نجاست» را معرفي كني.
      2. نمي‌خواهم درباره «اهل» يا درباره «كتاب» چيزي پيدا كنم.
      3. دنبال حتي مطلق «اهل كتاب» نيستم.
      4. من دنبال منابعي هستم كه درباره «نجاست اهل كتاب»‌ باشد.
2. در سطحي بالاتر:
   1. اصلاً بعضي از كتاب‌ها يك «باب» ويژه در اين موضوع دارند.
      1. ابواب اختصاصي اين موضوع را معرفي كن.
      2. نام باب را بياور و آدرس آن را در كتاب.
      3. مثلاً كتاب «وسائل الشيعه» يك باب اختصاصي درباره نجاست اهل كتاب دارد.
   2. حذف اطلاعات غيرمرتبط و ناكارآمد. اين‌ها بي‌دليل حجم نتايج را بالا مي‌برد.
      1. شناسنامه كتاب كه نبايد در جستجو بيايد.
      2. فهرست كتاب هم نبايد بيايد. زيرا رهزن است. تمام عناوين را دارد، ولي اطلاعاتي در اختيارم نمي‌گذارد.
      3. مقدمه‌اي كه ناشر نوشته، يا پيشگفتاري كه كتاب را معرّفي مي‌كند.
      4. خيلي از كتاب‌ها حجم زيادي پاورقي دارند كه موضوع اصلي جستجوي من نيستند.
3. در بالاترين سطح:
   1. «نجاست» با «طهارت» مرتبط است. مطالب مرتبط هم براي من مهم است.
      1. اگر جايي درباره «طهارت اهل كتاب» بحث شده است مي‌خواهم.
      2. اگر درباره «طهارت مشركان» هم بحث شده مهم است. زيرا بعضي اهل كتاب را جزو مشركان دانسته‌اند.
      3. درباره «نجاست مشركان» هم البته.
      4. اصلاً نه «نجاست» و نه «طهارت»، اگر در منبعي «مؤاكله اهل كتاب» هم آمده است مورد نياز من است. يعني هم‌غذا شدن با اهل كتاب. زيرا با نجاست و طهارت آن‌ها مرتبط است.
      5. فراتر از «مؤاكله»، هر منبعي كه بتواند چيزي از نجاست يا طهارت آن‌ها بفهماند، مانند «ازدواج با اهل كتاب». زيرا از آن نيز مي‌توانم حكم نجاست يا طهارت را پيدا كنم.
   2. نتايج را هم جمع‌بندي كن.
      1. هر كتابي از هر مؤلفي، در نهايت يك جمع‌بندي دارد. «جواهر الكلام» بالاخره قائل به نجاست اهل كتاب است يا طهارت آن‌ها.
      2. ده صفحه درباره اين موضوع نوشته است و در دو سه سطر آخر نظر نهايي داده.
      3. آن نظر را با يك نماد يا در قالب يك جمله كوتاه اطلاع بده.
   3. آمار هم نياز دارم.
      1. چند كتاب نظر مثبت دارند و چند كتاب نظر منفي.
      2. چند فقيه؟ نظر مثبت و منفي آن‌ها در موضوعي كه جستجو كرده‌ام.
      3. قرون و اعصار هم مهم است؛ چند نفر از قدما و چند نفر از متأخرين (فقهاي معاصر).
      4. اگر به شكل نمودارهاي دايره‌اي (Pie chart) هم نمايش داده شود كه چه بهتر.
   4. تفكيك ادله هم خيلي مفيد است.
      1. برخي براي حكم فقهي به آيات قرآن استناد كرده‌اند. اين آيات و استناد به آن‌ها در كتاب‌ها به صورت جدا ذيل بخش «ادله كتابي (قرآني)» قرار بگيرد.
      2. برخي به روايات استدلال كرده‌اند. آن بخش از كتاب‌ها در ذيل اين دسته قرار بگيرند.
      3. «اجماع» هم دليلي است كه بعضي فقها آورده‌اند.
      4. بعضي نيز «دليل عقلي» دارند. چهار نوع دليل داريم در تفقه، پس چهار دسته مي‌تواند منابع را تفكيك كند كه بسيار راهگشاست.

# انتظارات در پايان فاز نخست بازسازي

همه آن‏چه مورد انتظار پژوهشگر است را قطعاً نمي‌توانيم در فاز نخست تأمين كنيم. در اين مرحله هدف فقط اصطلاحاً «تميزكاري ديتاهاي موجود»‌ است. مي‌خواهيم داده‌هاي فعلي را به نحوي بازسازي كنيم كه حداقل انتظارات را فراهم سازد.

1. افزايش اطلاعات
   1. بعضي داده‌ها اضافه شده است به منابع قبلي كه بايستي خزش شوند.
      1. مدرسه فقاهت
         1. شيعه
         2. سني
      2. مدارك اسلامي
   2. دو منبع جديد بايستي به جُهد اضافه شوند.
      1. ويكي فقه
      2. ويكي شيعه
2. ارتقاي اطلاعات
   1. حذف بخش‌هاي غيرمفيد از جستجو
      1. شناسنامه
      2. صفحات خالي
      3. فهرست
      4. پاورقي
   2. شناسايي كتاب‌هاي تكراري در منابع مختلف
   3. تفكيك ماشيني محتواي كتاب
      1. سرصفحه‌ها و پاصفحه‌ها
      2. پاورقي‌ها
      3. علائم نگارشي و اصلاح پاراگراف‌بندي‌ها
   4. جايگزيني‌هاي ممكن
      1. حروفي مانند: «ي» و «ك» يكپارچه شود؛ يعني فقط يك حالت در بانك باشد.
      2. بعضي ويرايش‌ها به صورت ماشيني قابل انجام است:
         1. تبديل «كتاب ها» به «كتاب‌ها» يا «كتاب‌خانه» به «كتابخانه»
         2. براي اين‌كار خوب است يك ابزار جانبي در اختيار كارشناس باشد تا «از-به»ها را ثبت كند و سامانه به صورت خودكار جايگزيني را هنگام خزش انجام دهد.
         3. نسخه ناويراسته هم بايستي در بانك اطلاعات حفظ و نگهداري شود.
3. استفاده از عناوين معيار
   1. هر كتاب يك عنوان معيار دارد
      1. منابع از عناوين متفاوت براي كتاب‌ها استفاده مي‌كنند؛ «انتصار» به جاي «الانتصار في انفرادات الاماميه»
      2. يك جدول مستقل براي عناوين معيار كتاب‌ها
      3. بعدتر يك پژوهشگر مي‌تواند به صورت دستي اين موارد را يكپارچه كند
      4. هر كتاب يك اولويّت هم بتواند داشته باشد
         1. اين اولويّت پس از تكميل توسط پژوهشگر، مي‌تواند در نمايش نتايج استفاده شود.
         2. كتاب‌هاي با اولويّت بالاتر در نتايج جستجو بالاتر ديده خواهند شد.
   2. هر مؤلف هم عنوان معيار دارد
      1. بعضي منابع نوشته‌اند: «شريف مرتضي» به جاي اين‌كه بنويسند: «سيّد مرتضي علم الهدي»
      2. يك جدول مستقل براي عناوين معيار اشخاص
      3. اولويّت هم بشود براي اشخاص تعيين كرد تا در جستجو استفاده شود.
4. ايجاد صفحات معرفي كتاب و مؤلف
   1. با استفاده از ويكي‌هاي نامبرده
   2. كتاب‌ها و اشخاص بر اساس عناوين معيار به ويكي متصل شوند
   3. اطلاعات ويكي درباره آن‌ها در قالب صفحات شناسنامه نمايش داده شود

# اميدها براي آينده

فاز نخست بيشتر روي جمع‌آوري درست و صحيح اطلاعات تمركز دارد. ولي در آينده طرح‌هايي نياز است براي دستيابي به سطوح بالاتر ارائه خدمات. يكي از آن‌ها بازطراحي نماي سامانه و رابط كاربري، متناسب با امكانات و قابليت‌هاي جديدي‌ست كه داده‌ها ارائه خواهند كرد.

مي‌توانيم در ارتقاهاي بعدي سامانه براي ارائه خدمات در سطوح «هوشمند» مانند: تفكيك منابع بر اساس ادله فقهي، از كاربر هزينه‌اي نيز مطالبه كنيم و برنامه‌اي براي درآمدزايي داشته باشيم. چه بسا بتوانيم با دريافت هزينه، يك مقاله كامل به كاربر ارائه كنيم كه شامل تمامي فيش‌هاي اطلاعاتي و مستندات لازم به صورت صفحه‌بندي شده، همراه با پاورقي و عناوين فصول باشد!

سيدمهدي موسوي موشَّح - قم المقدّسه